DOCUMENT RESUME

ED 111 029 CS 202 250

AUTHOR	Butler, Patricia A.
TITLE	Evaluation of the First-Grade Spelling Component.
INSTITUTION	Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	SWRL-TM-3-71-08
PUB DATE	71
NOTE	32p.
EDRS PRICE	MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS	Communication Skills; Grade 1; *Language Skills; Primary Education; *Reading Instruction; *Spelling; *Spelling Instruction; Teaching Methods; *Word Study Skills

ABSTRACT

A component tryout of first-grade spelling was conducted using students who received concurrent reading instruction based on Southwest Regional Laboratory (SWRL) materials. Pretest and posttest measures obtained from four comparable first-grade classes were used as baseline data against which to judge program effectiveness. This report describes the results of the tryout of the first-grade spelling component. In most instances, weekly performance in tryout classes approximated or exceeded the 80 percent level. Posttest scores for tryout classes averaged 62 percent, 17 percent higher than the average of the comparison classes. The decline in performance from weekly test levels coupled with differences in performance on practice and transfer words suggests that the program's word attack procedures should be strengthened. Specific recommendations for improving these procedures are described, and suggestions are provided for modifying content, the word sequence, and teacher materials. (Author/LL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 111029

SOUTHWEST REGIONAL LABORATORY
TECHNICAL MEMORANDUM

DATE: August 20, 1971

NO: TM 3-71-08

TITLE: EVALUATION OF THE FIRST-GRADE SPELLING COMPONENT

AUTHOR: Patricia A. Butler

ABSTRACT

A component tryout of first-grade spelling was conducted using students who received concurrent reading instruction based on SWRL materials. Pre- and posttest measures obtained from four comparable first-grade classes were used as baseline data against which to judge program effectiveness. This paper reports the results of comparisons between tryout classes and classes that did not receive SWRL spelling instruction. Weekly test results obtained from tryout classes are also summarized.

US 802 250

This document is intended for internal staff distribution and use. Permission to reprint or quote from this working document, wholly or in part, should be obtained from SWRL, 11300 La Cienega Boulevard, Inglewood, California, 90304.

EVALUATION OF THE FIRST-GRADE SPELLING COMPONENT

It has often been pointed out that the processes of reading and spelling differ sufficiently along a number of dimensions to warrant separate, though not necessarily completely independent instruction (Cronnell, 1970d; Hanna & Moore, 1961; Hodges, 1965). Nonetheless as Hanna and Hanna (1965) have noted, spelling and reading skills may be considered mutually reinforcing. This would seem particularly true in instances in which instruction places primary emphasis on the learning of sound-to-letter or letter-to-sound associations that have high predictive utility and when explicit treatment is given to word analysis skills which make use of these high utility associations.

The findings of Cronnell (1970a) and of Shoemaker and Okada (1970) show that a substantial amount of transfer from reading instruction to spelling can be expected when these letter/sound conditions are met and when the overlap between tested and studied material is well controlled. The subjects in both of these experiments were kindergarten and first-grade children participating in the SWRL Communication Skills Program. The SWRL reading program includes a set of well defined word attack procedures which are used in teaching children to decode words into an initial consonant element and either a Vowel-Consonant (VC) or a Vowel-Consonant-Consonant (VCC) terminal element.

In the Cronnell study, kindergarten children spelled approximately 40% of their reading program words correctly although they had received no systematic spelling instruction during the school year. First-graders

in the Shoemaker and Okada study spelled approximately 52% of their program words correctly. Although these figures compare quite favorably with the results of more traditional transfer tests in which learned stimulus-response relations are reversed (cf. Jantz & Underwood, 1958), test scores fell far short of mastery. Although the SWRL reading program contributes to spelling proficiency, it is clear that a comprehensive communication skills program must include formal spelling instruction. The 1970-1971 component tryout of the First-Grade Spelling Program was designed to determine pacing and vocabulary load parameters for beginning spelling and to investigate the effectiveness of a number of activities which appeared appropriate for that level of instruction.

OVERVIEW OF THE PROGRAM

The general objectives of the spelling program parallel those of the reading program in that both emphasize the learning of generalizeable skills as well as the learning of specific elements and words. The program is composed of 190 study items divided into 18 study lessons. Study items include 71 word elements, 142 regularly spelled words and 48 sight words.

The program also includes two sets of transfer words. The first set consists of 72 words used in practicing word analysis skills. The second set consists of 36 words reserved for testing the level of skill development. With the exception of the 36 test-only transfer words, all items appear in the first- or second-grade reading materials. A more complete description of the program content appears in TN-3-71-2.

The activities designed to teach the lesson content and skills are based on a five-day work sequence which is briefly outlined in the Instruc-

tional Activities section of this paper. The material provided by SWRL includes study lists which categorized items as "sounds" (i.e., elements), "tricky words" (i.e., irregular or sight words) and "words" (i.e., regularly spelled words). Written exercises for students and sets of test dictation sentences for teacher use are also provided. Detailed information regarding program content, activities and materials is given in TN-3-71-2.

PROCEDURES EMPLOYED IN THE TRYOUT

At the beginning of the Spring semester of 1971, a pre-test was administered at two schools in the Torrance Unified School District and at one school in the Culver City Unified School District. The participants were enrolled in eight first-grade classes that were receiving concurrent reading instruction in the SWRL Communication Skills Program. The pre-test consisted of 32 items, with eight items sampled from each of the program's four outcome categories (elements, regular words, transfer words, and sight words). The contents of the pre-test are listed in the appendix section of TN-3-71-06.

Tryout and comparison groups were established by pairing classes on the basis of mean pre-test scores. Because of the possibility that tryout teachers might inadvertently share specific program information with comparison teachers, an effort was made to pair classes which were drawn from different schools. Because of the distribution of classes over schools, this restriction held for only three of the four pairs of classes.

Instruction in tryout classes was initiated at the end of January and continued until the first week of June. During the second week of

June, a 32 item posttest was administered to tryout and comparison classes. Although pre-test and posttest items were different, the two tests were identical in form. Posttest content is listed in TN-3-71-07.

INSTRUCTIONAL ACTIVITIES

Tryout Classes. In tryout classes, instructional procedures conformed in a general sense to the guidelines described in TN-3-71-2. In each class, weekly lessons began with two practice days, with from 5 to 30 minutes devoted to oral drills[1] and an average of 20 minutes devoted to written exercises. The third day of each week was devoted to a practice test which included all study items and all word-attack practice items. The fourth day was devoted to a 12 item final test which included 10 study items and two transfer items. The fifth day was reserved for review and remediation.

Although departures from suggested procedures occurred on a number of lessons, the modified procedures were generally consistent with the program's objectives. Only two major exceptions were noted. First, in some instances, teachers adopted practice testing procedures which did not provide students with immediate inter-item feedback. Second, and perhaps more important, teachers failed to reinforce the program's word-attack strategy when children had difficulty spelling regular words during oral drills. A detailed account of procedural variations based on classroom observation is included in TN-3-71-06.

---

[1] The recommended procedures for oral drills never required more than 10 minutes. The higher figures occurred in classes in which teachers extended or modified suggested activities.

On the Monday following the completion of every fourth or fifth study lesson, a 16 item comprehensive test reviewing all preceding material was administered by SWRL staff. Later that day or on the following day, a week-long review lesson was conducted by the teacher. Review material for each class was selected by the teacher using error data from the preceding four or five weekly tests. At the time of posttest administration, tryout classes had completed 17, 18, 19 and 21 lessons respectively.

Comparison Classes. No observational data are available for the comparison classes. Information compiled at the end of the school year indicated that all classes used the state adopted text, Basic Goals in Spelling. At the first-grade level, most words in the state text are treated as sight words. Only word initial consonant sound-to-letter correspondences are taught.[2] In addition to using the state text, one class also received practice on a set of high frequency sight words selected from Success in Spelling.

The amount of time devoted to instruction each week ranged from 30 minutes to an hour in the four comparison classes. This time was distributed over from one to five daily sessions. By the second week of June, one comparison class had completed approximately two-thirds of the material in the state series first-grade workbook. The other classes completed the entire book. The class that received additional material completed instruction on 96 sight words as well.

---

2 The Basic Goals in Spelling series is reviewed by Cronnell in TM-2-71-22.

PERFORMANCE OF THE TRYOUT POPULATION DURING THE COURSE OF INSTRUCTION

Weekly Test Performance. Weekly test results averaged over classes are reported in Tables 1 and 2. Individual class ranges on the first 18 lessons are reported in Table 3. Weekly scores of each participating class are reported in TN-3-71-06 and TN-3-71-07. Because no fixed number of elements, regular words, or sight words was set for end-of-week tests, the distribution of study items differed to some extent over tests. Therefore, no distinction is made among outcomes in Table 1. Each test included at least one item from each of the categories represented in the study material.

The percentages in Table 1 indicate that when students in both reading programs are considered, weekly test scores on study items averaged between 77.4% and 91.6%. The performance of students in the Second-Year Communication Skills Program (SYCSP) was generally quite good, ranging from 82.7% to 95.7% over weeks. The performance of students in the Transition Program (TP) was both poorer and more variable; scores ranged between 56.7% and 77.3% over tests.

Transfer scores, which are shown in Table 2, were generally lower than study item scores on each test. In fact, Schwab has noted in TN-2-71-21, that the overall error rate of transfer words was approximately 14% higher than that of study words. For the tryout group as a whole, the level of transfer performance varied from 50.6% to 80.2% over tests. SYCSP students spelled from 59.3% to 87.1% of the transfer items correctly on weekly tests. TP students spelled 22.2% to 60.9% of the words correctly.

TABLE 1

Tryout Population Scores for Weekly Study Items[1]: Mean Percent Correct as a Function of Reading Level

Lesson:	TP	SYCSP	Both
1	77.3	95.7	91.6
2	65.0	91.4	85.2
3	62.6	89.8	85.6
4	64.5	90.8	85.5
6	60.9	90.9	84.0
7	67.6	83.1	80.0
8	73.9	93.8	89.4
9	65.2	93.3	86.5
10	70.0	90.9	86.5
12	56.8	82.7	77.4
13	63.0	86.9	81.6
14	62.0	90.0	83.4
15	56.2	87.5	81.4
17	74.0	91.7	87.6
18	56.7	85.1	78.4

[1] Lessons 5, 11, and 16 were omitted because they were reviews composed of different items in each class. Lessons 19-21 are omitted because they were completed by fewer than two classes.

TABLE 2

Tryout Population Scores for Weekly Transfer Items[1]: Mean Percent Correct as a Function of Reading Level

Lesson:	TP	SYCSP	Both
2	57.69	87.06	80.18
3	43.48	77.71	70.28
4	59.09	85.47	80.09
6	33.93	69.87	63.37
7	50.00	71.43	67.14
8	60.87	81.48	76.92
9	50.00	81.01	73.56
10	46.51	78.16	70.91
12	29.53	60.71	54.25
13	36.96	77.16	68.27
14	58.00	79.63	74.53
15	45.24	81.93	74.77
17	44.00	80.25	71.70
18	22.22	59.32	50.65

[1] Lessons 5, 11 and 16 have been omitted because they consisted of reviews which were composed of different items in each class. Lesson 1 is omitted because it included no transfer words. Lessons 19-21 are omitted because they were completed by fewer than two classes.

Although the averages reported in Tables 1 and 2 are quite high in most instances, it is important to note that the performance of individual classes was quite variable. The ranges reported in Table 3 illustrate this point. Only the students in one class, T-1, achieved consistently high levels of performance. Performance in other classes was less stable, particularly for TP students, whose weekly test scores varied by as much as 47% in one class and 33% in another.

Comprehensive Test Performance. As noted above, three 16 item comprehensive tests were administered by SWRL staff at four to five week intervals. The content of each test was sampled from the entire population of items covered in preceding lessons. Thus, the first test reviewed material from lessons 1 through 4; the second test covered lessons 1 through 10, and the third test covered lessons 1 through 15. The content of these tests appears in TN-3-71-06 and TN-2-71-07.

The samples used in each test reflected the relative frequency with which each outcome occurred over the series of completed lessons. Thus, the first review test included no sight words. This sampling procedure was deemed more appropriate to determining retention of studied material than procedures which would weight all outcomes equally, regardless of their representation in the total word pool.

Comprehensive test results are presented along with pre- and post-test data in Table 4. The two most outstanding features of these data are the consistent superiority of SYCSP scores and the overall decline in performance over the last three tests. This decline is not surprising in view of the fact that each test included material of increasing diffi-

TABLE 3

Range of Percent Correct Scores of Each Tryout Class on the First 18 Lessons.[1]

Class		TP	SYCSP	Average
T-1	N[2]	4-6	14-18	20-24
	Study	82.0-100.0	94.4-100.0	91.8-100.0
	Transfer[3]	20.0-100.0	65.6-100.0	54.8- 97.5
		(67.0-100.0)	(82.1-100.0)	(82.0-100.0)
T-2	N		20-32	
	Study		71.3- 91.1	
	Transfer		56.0- 87.1	
T-3	N	4-8	20-23	26-30
	Study	51.4- 98.0	80.0- 99.6	77.4- 96.3
	Transfer[3]	25.0-100.0	38.6- 95.4	37.0- 90.4
			(65.9- 95.4)	(60.3- 90.4)
T-4	N	10-14	14-18	25-32
	Study	35.0- 68.4	75.7- 95.6	59.6- 78.3
	Transfer[3]	4.6- 58.3	56.7- 88.2	38.5- 70.7
		(22.7- 58.3)		(50.0- 70.7)

[1]The range for class T-3 is based on only 17 lessons.

[2]N indicates the number of students present on various test days.

[3]Two lessons, 12 and 18, introduced new concepts, long vowels and suffixes. Transfer scores on these lessons were often much lower than on other lessons. Where exclusion of these scores resulted in a much narrower range, the more representative figures are shown in parentheses.

TABLE 4

Mean Percent Correct Scores for Tryout Population on Each of Five Tests.

		N	Elements	Regular Words	Transfer Words	Sight Words	Total Score
PRETEST	TP	29	33.62	25.86	14.22	21.98	23.92
	SYCSP	81	45.22	32.56	27.78	46.45	38.00
	AVERAGE	110	41.48	30.34	24.20	41.02	34.29
REVIEW TEST 1	TP	23	68.12	52.26	47.83		58.15
	SYCSP	81	89.51	85.96	82.10		86.81
	AVERAGE	104	84.78	78.73	74.52		80.47
REVIEW TEST 2	TP	24	54.17	55.73	29.17	50.00	51.56
	SYCSP	89	81.35	82.30	63.48	82.02	79.63
	AVERAGE	113	75.58	76.66	56.19	75.22	73.67
REVIEW TEST 3	TP	25	44.00	51.50	42.00	36.00	46.50
	SYCSP	85	65.29	81.03	65.29	70.00	73.75
	AVERAGE	110	60.45	74.32	60.00	62.27	67.57
POSTTEST	TP	24	56.25	49.48	40.63	54.17	50.13
	SYCSP	81	74.69	64.20	56.79	64.97	65.16
	AVERAGE	105	70.48	60.83	53.10	62.50	61.73

culty. This increased difficulty resulted from increases in the proportion of irregularly spelled words over successive blocks of material and increases in the a priori difficulty level of vowel correspondences over blocks.[3]

The cell entries in Table 4 indicate that the two trends noted above were evident on individual outcome scores as well as on total scores. It is clear, however, that not all outcomes were equally affected by performance declines. Although the pattern changes depending on which test and which reading level is selected, transfer scores appeared to be less stable in general.

Error Analysis. Schwab (1971) has analyzed specific errors for all words included in end-of-week tests. The results of this analysis are reported in TN-2-71-21. In brief, her findings were that errors in the spelling of individual correspondences occurred at a rate of 13% and that these errors fell into several broad catergories which may have implications for instruction.

Among the categories isolated by Schwab are those errors which probably reflect sound discrimination difficulties and which may be corrected by the use of drills which provide practice in distinguishing minimal pairs such as fat/vat, bag/beg, bat/bad, and sad/sand. Another error source may possibly be traced to the existence of many social and regional dialects which do not distinguish certain sounds (as in the minimal pairs which/witch, ten,/tin) which are distinguished in other dialects. Because of dialect variations of this type, certain rule-based words may have to

[3]A priori difficulty levels were based on error data reported by Cronnell in TN-2-70-12. These data were compiled from responses of second-grade students who had received SYCSP reading instruction but no formal spelling instruction.

be treated as sight words and learned by memorization.

Other errors resulted from the fact that many sounds have several admissible spellings. Where the correct option is predictable given certain environmental constraints, spelling proficiency should be improved by instruction which explicitly recognizes the dependence of spelling on context. Where low predictability is involved, optional patterns will have to be memorized in individual sight words or in lists that contain words that possess the same irregular feature.

Schwab also noted a number of errors which seemed attributable to the fact that members of long and short vowel minimal pairs (e.g., h<u>i</u>d<u>e</u>/h<u>i</u>d) were taught in a single lesson without sufficient contrastive emphasis being placed on either the aural or the graphic characteristics of the words, or on the relationship between the two sets of characteristics. Undoubtedly, performance should be improved by proper contrast exercises.

PRE- AND POSTTEST PERFORMANCE OF TRYOUT AND COMPARISON POPULATIONS

Total scores for individual tryout and comparison classes are shown in Table 5. Entries in Table 5 are labeled to indicate pairing of tryout and comparison classes. The letters T and C denote tryout and comparison classes, respectively. The digit in each letter-number code indicates the rank assigned to a particular class on the basis of total pre-test scores. Because ranks were used in matching classes, labels also indicate which tryout and comparison classes constitute a pair. Thus, the labels T-1 and C-1, for example, refer to the tryout and comparison classes that achieved the highest level of pre-test performance.

The total scores in Table 5 show that on the whole, the two groups were fairly well matched. The comparison group averaged 35.23% correct

TABLE 5

Pre- and Posttest Scores: Mean Percent Correct
for Each Tryout and Comparison Class

Class		TP	SYCSP	Average
T-1	PRE	40.06 (N=11)	57.10 (N=11)	47.87 (N=22)
	POST	81.25 (N= 8)	89.38 (N=10)	85.76 (N=18)
T-2	PRE		35.21 (N=30)	
	POST		50.58 (N=27)	
T-3	PRE	10.55 (N= 8)	40.08 (N=23)	32.26 (N=31)
	POST	33.04 (N= 7)	71.45 (N=22)	62.18 (N=29)
T-4	PRE	16.88 (N=10)	27.57 (N=17)	24.42 (N=27)
	POST	35.76 (N= 9)	75.00 (N=15)	60.29 (N=24)
ALL TRYOUT CLASSES	PRE	23.92 (N=29)	38.00 (N=81)	34.29 (N=110)
	POST	50.13 (N=24)	65.16 (N=81)	61.73 (N=105)
C-1	PRE	34.38 (N= 5)	71.53 (N=18)	63.45 (N=23)
	POST	31.25 (N= 5)	59.20 (N=18)	53.13 (N=23)
C-2	PRE		31.68 (N=22)	
	POST		35.59 (N=18)	
C-3	PRE	26.10 (N=17)	41.67 (N= 6)	30.16 (N=23)
	POST	29.46 (N=14)	55.00 (N= 8)	36.18 (N=22)
C-4	PRE	23.44 (N=22)	78.13 (N= 1)	26.22 (N=23)
	POST	40.78 (N=20)	68.75 (N= 1)	42.11 (N=21)
ALL COMPARISON CLASSES	PRE	25.36 (N=44)	44.48 (N=47)	35.23 (N=91)
	POST	35.90 (N=39)	54.09 (N=42)	45.33 (N=81)

on the pre-test and the tryout group averaged 34.29% correct. Although some individual pairings of classes were clearly better than others, as Table 5 shows, in two cases the tryout class had the higher score and in the other two cases, the comparison group did.

Table 5 also shows better end-of-year performance for the tryout students. Class scores ranged between 50% and 85.76% in the tryout group and between 35.59% and 63.45% in the comparison group. In spite of this overlap of scores, each tryout class achieved a higher score than its matched reference class. The overlap is attributable primarily to the fact that SYCSP students in one comparison class outperformed SYCSP students in three of the four tryout classes, while one tryout class was outperformed by three of the four comparison classes.

STATISTICAL ASSESSMENT OF TRYOUT AND COMPARISON GROUP DIFFERENCES

Design and Analysis. Tryout results were tested using a mixed analysis of variance design (Myers, 1967). The basic cell in this design is defined by combinations of the following four variables: Group[4] (Tryout/Comparison), Reading Program (TP/SYCSP), Test (Pre-test/Posttest), and Outcome (Elements/Regular Words/Transfer Words/Sight Words).

The Sample. Although it would have been desirable to analyze the data of all students for whom pre-test and posttest results were available, it was necessary to work with only a sample. Sampling was necessitated by the fact that the distribution and number of TP and SYCSP students differed in tryout and comparison groups.

The sample consisted of 96 students, with 24 students representing

[4]Because one comparison class and one tryout class consisted entirely of SYCSP students, the distinction among classes within groups was ignored in order to avoid an unbalanced design.

each of the four combinations of Group and Reading Program. Sampling was random, with the single restriction that within both groups, the number of students selected from any class be proportional to the number of students in that class included in the population.

In order to guard against the possibility that the results of this analysis were unduly influenced by program modifications introduced by any particular teacher, a second analysis was run using sample data in which all classes were weighted equally. Although means and specific values of F differed for the two differently constituted samples, the pattern of results in the two analyses is essentially the same, therefore only the results of the first analysis are reported.

Results. Results of the analysis of variance are summarized in Table 6. Although Table 6 contains several additional terms, only those sources for which the effect of Groups (G) and Tests (T) are partialled out are critical to evaluation of program effectiveness. Therefore, other effects which achieved statistical significance are reported with little additional comment.

As expected, SYCSP students scored higher than TP students and overall posttest performance surpassed pre-test performance. These results are reflected by significant R and T main effects noted in Table 6.

The overall effect of outcomes was also significant, with element scores being highest and transfer scores lowest. Although this pattern held on the pre-test and the posttest, the interaction of Tests and Outcomes was significant. Mean outcome scores are presented in Table 7. As the marginal entries in Table 7 suggest, the TO interaction is probably

TABLE 6

Analysis of Variance of Sample Pretest and Posttest Scores.

Source of Variation	Sum of Squares	df	Mean Square	F
Total	4602.17	767		
Between S	2617.78	95	27.56	
G (Group)	44.56	1	44.56	1.97
R (Reading Program)	495.69	1	495.69	21.96**
GR	.95	1	.95	<1.00
S/GR	2076.58	92	22.57	
Within	1984.39	672		
T (Test)	457.82	1	457.82	115.61**
GT	156.78	1	156.78	39.59**
RT	7.32	1	7.32	1.85
GRT	3.13	1	3.13	<1.00
ST/GR	364.33	92	3.96	
O (Outcome)	178.12	3	59.38	41.52**
GO	7.10	3	2.36	1.65
RO	3.68	3	1.23	<1.00
GRO	2.07	3	.69	<1.00
SO/GR	395.15	276	1.43	
TO	12.13	3	4.04	2.97*
GTO	5.23	3	1.74	1.27
RTO	13.98	3	4.66	3.43*
GRTO	.92	3	.31	<1.00
STO/GR	376.63	276	1.36	

* $p < .05$
** $p < .001$

TABLE 7

Pre- and Posttest Sample Means and Mean Percentages[1] of TP and SYCSP Students on Each of the Four Spelling Outcomes.

	PRETEST				POSTTEST			
	Elements	Regular Words	Transfer Words	Sight Words	Elements	Regular Words	Transfer Words	Sight Words
TP	2.75 (34.38)	2.10 (26.25)	1.28 (16.00)	1.90 (23.75)	3.69 (46.13)	3.32 (41.50)	2.79 (34.88)	3.88 (48.50)
SYCSP	4.27 (53.38)	2.86 (35.75)	2.71 (33.88)	4.04 (51.00)	5.60 (70.00)	5.29 (66.13)	4.54 (56.75)	5.44 (68.00)
AVERAGE	3.50 (43.75)	2.48 (31.00)	1.99 (24.88)	2.97 (37.13)	4.64 (58.00)	4.30 (53.75)	3.67 (45.88)	4.66 (58.25)

[1] Percentages are shown in parentheses.

attributable to the fact that pre-test scores were much more variable than posttest scores. The cell entries show that for SYCSP students, the largest pre- to posttest gain occurred on regular words, while for TP students, the largest gain occurred on sight words. SYCSP and TP gains on the remaining outcomes were more nearly equivalent. These differences in the pattern of change over outcomes resulted in a significant RTO interaction.

Contrary to expectations, the average performance of the tryout group did not differ appreciably from that of the comparison group. The means in Table 8 suggest a plausible explanation for this finding. It appears that the initial superiority of the comparison students and the posttest superiority of the tryout students combined to cancel out any overall effect of Group.

The averages shown in Table 8 indicate that end-of-year performance represented an improvement over earlier performance in both tryout and comparison groups. In the tryout group, the average outcome score on the pretest was 2.49 out of a possible 8 points and 5.01 on the posttest. In the comparison group, the increase was smaller, from 2.98 to 3.63. Posttest scores converted to percentages were 62.59 for the tryout group and 45.33 for the comparison group.

The divergence of group scores over tests resulted in a significant GT interaction. Simple effects tests indicated that while pre-test differences were not statistically reliable, posttest differences were ($\underline{F}[1,92] = 11.52$, $p < .01$). It may be concluded therefore, that the program enhanced spelling proficiency over and above the level which may be attributed to any transfer from SWRL reading instruction.

TABLE 8

Sample Means and Mean Percentage Scores for Tryout and Comparison Groups Averaged Over Outcomes.[1]

		Pretest	Posttest
Tryout Group	TP	1.76 (22.00)	4.01 (50.13)
	SYCSP	3.22 (40.22)	6.00 (75.06)
	Average	2.49 (31.11)	5.01 (62.59)
Comparison Group	TP	2.24 (28.09)	2.82 (35.31)
	SYCSP	3.71 (46.47)	4.43 (55.34)
	Average	2.98 (37.28)	3.63 (45.33)

[1] Percentages given in parentheses.

The cell means shown in Table 8 indicate that instruction had no differential effects on students in the two reading programs. As these data show, TP and SYCSP scores within the tryout group increased by amounts which were roughly similar. The same is true for scores within the comparison group. However, as noted above, increases in the comparison group were of smaller magnitude.

The data failed to reveal any evidence that instruction selectively affected performance on the four outcomes. The mean percent scores in Table 9 show that in the tryout group, the range of scores was fairly stable over tests, with all outcomes exhibiting increases close to 30% from the pretest to the posttest. In contrast, increments in the comparison group varied considerably as a function of outcome; elements showed practically no change over tests, while transfer words, which showed the greatest amount of change, increased by approximately 13%.

DISCUSSION

The results of the tryout indicated that the materials and procedures used in the first-grade spelling tryout are capable of augmenting spelling proficiency. However, the 62% posttest performance suggests that there is room to improve the program effectiveness. Performance levels of this proportion are not uncommon for initial tryouts. Sullivan (1968) reported end of year performance of 63% in the 1967-68 tryout of the First-Year Communication Skills Program. Niedermeyer (1970) reported midyear performance of 73% in the 1968-69 tryout of the revised version of that program.

Although in this context, overall performance may be considered satis-

TABLE 9

Pre- and Posttest Sample Means and Mean Percentages[1] of Tryout and Comparison Groups on Each of Four Outcomes

Tryout Group	Elements	Regular Words	Transfer Words	Sight Words
Pretest	3.31 (41.38)	2.29 (29.63)	1.90 (23.69)	2.46 (30.75)
Posttest	5.56 (69.50)	5.02 (62.75	4.22 (52.81)	5.22 (65.31)
Comparison Group				
Pretest	3.70 (46.31)	2.74 (33.31)	2.08 (26.06)	3.48 (43.44)
Posttest	3.73 (46.63)	3.58 (44.81)	3.10 (38.81)	4.08 (51.06)

[1] Percentages are shown in parentheses.

factory, the wide variability among tryout classes is cause for concern. As noted earlier, the performance of one tryout class fell below that of three out of four comparison classes. Because all tryout classes were exposed to the same content and materials, differences in instructional procedures may provide the only plausible explanation of this finding.

The possibility that instructional procedures are the critical factor seems reasonable in view of the fact that the class with the poorest performance record received only the amount of oral practice recommended in the Teachers' Guide. In all other classes, teachers viewed the prescribed amount of work as insufficient, and as a result, provided other appropriate oral exercises. Apparently, these additional adtivities contributed to greater learning, as reflected by end-of-week test performance, and to greater long term retention, as reflected by comprehensive and posttest performance.

It should be noted at this point that the primary purpose of oral drills was to provide students with teacher-directed practice in segmenting regular words into component sounds and in constructing the spelling of these words by correctly sequencing the letters which represent each sound. Although every weekly lesson included a written exercise which required students to generate four new words based on one or more of the elements in the current study list, classroom observations indicated that teachers did not often stress the analytic nature of the task. Thus, in a fair number of instances, the bulk of word attack practice occurred in the context of oral exercises.

Given this conception of the purpose and importance of oral drills,

another indication of the weakness of current procedures for conducting oral practice is the finding that transfer scores were typically lower than both element and regular word scores. Although it is true that transfer scores were deflated in some instances by an unfortunate choice of test items (e.g., the word "tin" which is not easily distinguished in a dictation test from the word "ten"), it is also true that discrepancies among element, regular word and transfer word performance should be quite small if children have in fact learned a general strategy for spelling new words which are composed of familiar elements. The finding of rather sizeable differences on both immediate and delayed tests suggests that word analysis skills are being taught with less than maximal effectiveness.

Although statistical comparisons revealed no difference in instructional effectiveness for students working in different reading programs, in view of differences in entry level performance, it is clear that students who begin with an advantage will maintain that advantage throughout the program. Therefore, if a single standard of performance is to be established for all students, it will be necessary to devise some method of increasing the learning rate of the less advanced students.

Because teachers indicated that the amount of material and the number of exercises used in the tryout provided more than enough work for slower readers and poorer spellers, no additional remedial material is planned for inclusion in the revised program. Instead, it is hoped that the use of better word attack procedures will reduce any advantage attributable to prior familiarity with specific items which appear in both the reading and spelling materials.

SUGGESTED REVISIONS

Revisions, other than minor editorial revisions, fall in the broad categories of content, sequencing and organization, word attack procedures, and teacher materials. These modifications are based on teacher comments, classroom observation, and evaluations of the program's effectiveness as reflected by measured performance.

Content. The content of the program used in the tryout was selected from the word and element lists of the Model 1 first- and second-year reading materials. Subsequent development will expand the content to include a greater number of high frequency sight words which appear on the Dolch and Rinsland lists and on standardized tests which are used locally.

High priority will be given to rule words which appear in the Model 2 reading word base. A primary advantage of this policy is that the tryout of revised materials with children using Model 1 readers can provide information regarding the program's effectiveness under conditions in which the spelling vocabulary is not completely determined by the content of the reading program. This type of information would be particularly useful in the event that the spelling and reading components of the Communication Skills Package are marketed separately.

Sequence and Organization. Material for the 1971 tryout was sequenced using a complex set of decision criteria. As a result of using these criteria, words are grouped according to common vowel correspondence and ordered according to increasing difficulty. Difficulty levels were established using error data for individual sound-to-letter correspondences

compiled by Cronnell (1970b). The results of the tryout suggested that a more optimal sequence could be generated if discriminability as well as difficulty is taken into account.

One characteristic of such a sequence would be that easily confused sounds would be introduced in separate, nonadjacent units. A second characteristic is that words would be grouped according to a scheme which would permit contrasts of the confusable sounds at that point in the sequence at which the second member of the confusable pair is introduced.

Word Attack Practice. Current word attack procedures have been formulated in terms which make them applicable to all lessons, but which are specific enough to describe exactly what children should be required to do in each oral drill. A typical instruction is "Read aloud three or four regularly spelled words and call on individual children to sound out and spell each without the aid of the study list". Among the difficulties associated with current procedures are that they provide a minimal amount of word attack practice, leave little room for variation, and require that the teacher determine which words and which situations deserve special attention.

The latter characteristic is critical, it would seem, in view of the fact that error patterns suggest that certain topics (e.g., long vowels, optional spellings of [k]) might have benefitted from special treatment. Revised word attack instruction should be worked out separately for each lesson, although an attempt should be made to identify several major drill formats which can be used with different content. Instruction should include both oral and written exercises which make use of minimal pairs,

and of rhyming and alliterative patterns. At least initially, exercises should be structured in such a way that the student is forced to use the word building strategy that the word attack model implies. Because of the high error rates associated with vowels, some exercises should focus on individual vowel correspondences, although children should be encouraged to build words by spelling elements as long as this approach is used in reading.

Teacher Materials. At the present time, teachers are given very few special instructions for individual lessons. When special instructions are provided, they are quite brief and are generally specific to a particular word (e.g., "The word 'use' is considered a verb in all exercises. Be sure that the children sound the letter s as 'z-z-z' in oral practice.") or to a particular exercise (e.g., "Note that in item 4 two letters, sh, replace a single letter.")

Revised materials should provide sufficient information to aid the teacher in preparing a well integrated, conceptually sound lesson each week. For each lesson, the teacher should be provided with relevant linguistic background information which includes both general principles and special information regarding dialect variations which may produce interference. When appropriate, teacher notes should also include information regarding common errors, particularly difficult words, rules and mnemonics which may be helpful to the children, and important word, element, or vowel correspondence contrasts.

SUMMARY

This report described the results of the tryout of the first-grade spelling component. In most instances, weekly performance in tryout classes approximated or exceeded the 80% level. Performance on periodically administered comprehensive tests was somewhat lower. Posttest scores for tryout classes averaged 62%, which although not an extremely impressive figure, is approximately 17% higher than the average of the comparison classes.

The decline in performance from weekly test levels coupled with differences in performance on practice and transfer words suggested that the program's word attack procedures should be strengthened. Specific recommendations for improving these procedures are described along with suggestions for modifying content, the word sequence, and teacher materials.

REFERENCES

Butler, Patricia A. Content and specifications for the Mod 1 1970 spelling program. Technical Note No. TN 3-71-2, 1971, Southwest Regional Laboratory, Inglewood, California.

Butler, Patricia A. End of year test results for the 1970-1971 first-grade spelling tryout. Technical Note No. TN-3-71-07, 1971, Southwest Regional Laboratory, Inglewood, California.

Butler, Patricia A. Interim report for the 1970-1971 first-grade spelling component tryout. Technical Note No. TN-3-71-06, Southwest Regional Laboratory, Inglewood, California.

Cronnell, Bruce. Analysis of spelling ability of participants in SYCSP with no formal spelling training. Technical Note No. TN-2-70-28, 1970a, Southwest Regional Laboratory, Inglewood, California.

Cronnell, Bruce. Analysis of spelling errors made by participants in SYCSP with no formal spelling training. Technical Memorandum No. TM 2-70 12, 1970b, Southwest Regional Laboratory, Inglewood, California.

Cronnell, Bruce. Beginning spelling: a review of six spelling series. Technical Memorandum No. TM-2-71-22, 1970c, Southwest Regional Laboratory, Inglewood, California.

Cronnell, Bruce. Reading and spelling. Technical Note No. TN-2-70-15, 1970d, Southwest Regional Laboratory, Inglewood, California.

Hanna, Paul R., and Hanna, Jean S. Applications of linguistics and psychological cues to the spelling course of study. Elementary English, 1965, 42, 753-759.

Hanna, Paul R., and Moore, James T. Spelling--from spoken word to written symbol. Elementary School Journal, 1953, 53, 329-337.

Hodges, Richard E. The psychological bases of spelling. Elementary English, 1965, 42, 629-635.

Jantz, Eleanore M., and Underwood, Benton J. R-S learning as a function of meaningfulness and degree of S-R learning. Journal of Experimental Psychology, 1958, 56, 174-179.

Myers, Jerome L. Fundamentals of Experimental Design, ch. 8, Boston: Allyn and Bacon, Inc., 1966.

Niedermeyer, Fred C. Developing exportable teacher training for criterion-referenced instructional programs. Technical Report No. TR 22, 1970, Southwest Regional Laboratory, Inglewood, California.

Schwab, Donna R. An analysis of spelling errors in the 1971 c tryout of a first-grade spelling component. Technical Note No. TN-2-71-21, 1971, Southwest Regional Laboratory, Inglewood, California.

Shoemaker, David M., and Okada, Masahito. The communication skills program and concomitant spelling proficiency. Technical Note No. TN-3-70-3, 1970, Southwest Regional Laboratory, Inglewood, California.

Sullivan, Howard S. 1967-68 classroom tryout of the SWRL first-year communication skills program. Unpublished report, 1968, Southwest Regional Laboratory, Inglewood, California.